



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

Microarray resources for genetic and genomic studies in chicken: a review

Citation for published version:

Gheyas, AA & Burt, DW 2013, 'Microarray resources for genetic and genomic studies in chicken: a review', *genesis: The Journal of Genetics and Development*, vol. 51, no. 5 (Sp Iss S1), pp. 337-356.
<https://doi.org/10.1002/dvg.22387>

Digital Object Identifier (DOI):

[10.1002/dvg.22387](https://doi.org/10.1002/dvg.22387)

Link:

[Link to publication record in Edinburgh Research Explorer](#)

Document Version:

Publisher's PDF, also known as Version of record

Published In:

genesis: The Journal of Genetics and Development

Publisher Rights Statement:

Available under Open Access

General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact openaccess@ed.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.



REVIEW

Microarray Resources for Genetic and Genomic Studies in Chicken: A Review

Almas A. Gheyas* and David W. Burt

Department of Genetics and Genomics, The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Midlothian, United Kingdom

Received 20 December 2012; Revised 21 February 2013; Accepted 22 February 2013

Summary: Advent of microarray technologies revolutionized the nature and scope of genetic and genomic research in human and other species by allowing massively parallel analysis of thousands of genomic sites. They have been used for diverse purposes such as for transcriptome analysis, CNV detection, SNP and CNV genotyping, studying DNA-protein interaction, and detection of genome methylation. Microarrays have also made invaluable contributions to research in chicken which is an important model organism for studying embryology, immunology, oncology, virology, evolution, genetics, and genomics and also for other avian species. Despite their huge contributions in life science research, the future of microarrays is now being questioned with the advent of massively parallel next generation sequencing (NGS) technologies, which promise to overcome some of the limitations of microarray platforms. In this article we review the various microarray resources developed for chicken and their past and potential future applications. We also discuss about the future of microarrays in the NGS era particularly in the context of livestock genetics. We argue that even though NGS promises some major advantages—in particular, offers the opportunity to discover novel elements in the genome—microarrays will continue to be major tools for research and practice in the field of livestock genetics/genomics due to their affordability, high throughput nature, mature established technologies and ease of application. Moreover, with advent of new microarray technologies like capture arrays, the NGS and microarrays are expected to complement each other in future research in life science. *genesis* 51:337–356. © 2013 Wiley Periodicals, Inc.

Key words: gene expression array; genotyping array; aCGH; capture array

INTRODUCTION

The research and applications in the field of genetics and genomics have seen tremendous progress over the last two decades. While it is the human genomics that has experienced the greatest development, other organisms—both animals and plants—are following suit. From among the farm animal species, chicken has emerged as an important model organism for research in diverse fields such as for studying embryological development, immunology, oncology, virology, evolution, genetics and genomics, and especially as a model for other avian species. For instance, the ready access to its embryos in incubated eggs and ease of manipulating the embryos for experimental purposes, have made chicken an ideal candidate for studying vertebrate embryonic developments (Brown *et al.*, 2003; Stern, 2005) and much of the insights on the human limb formation have originated from studying chicken (Schmutz and Grimwood, 2004). Chicken has made invaluable contributions to the field of immunology through its use for elucidating the roles of lymphocytes in adaptive immunity, discovery of B cells, studying the mechanisms of gene conversion in the unique environment of Bursa, discovering the roles of chicken Major Histocompatibility Complex (MHC) in rendering resistance to infectious diseases, and by becoming one of the first model systems for discovering vaccines against viral diseases

* Correspondence to: Almas A. Gheyas, Department of Genetics and Genomics, The Roslin Institute and Royal (Dick) School of Veterinary studies, University of Edinburgh, Easter Bush, Midlothian, EH25 9RG, Scotland, UK. E-mail: almas.ghayas@roslin.ed.ac.uk

Published online 7 March 2013 in

Wiley Online Library (wileyonlinelibrary.com).

DOI: 10.1002/dvg.22387

(Davison, 2003). It has also been used as a model for human genetic disorders such as muscular dystrophy, epilepsy and decreased immunological responses and chicken lines that show similar symptoms to those of human patients have been created for these diseases (Schmutz and Grimwood, 2004). Chicken has also been considered to be an important species for evolutionary and comparative genomic studies both for closely related and distant species. Since chicken has diverged from humans more than 310 million years ago, it is considered a good example of an "outgroup" which can be used to identify highly conserved genomic regions having important regulatory and functional roles (International Chicken Genome Sequencing Consortium, 2004). Due to its importance as a major farm animal, chicken has been a focus of diverse genetic and genomic research for many years leading to an accumulation of vast knowledge-base and resources for this species. For instance, chicken is the first farm animal to have its full genome sequenced (International Chicken Genome Sequencing Consortium, 2004). A high resolution consensus linkage map consisting of >9,000 SNP and microsatellite markers are now available showing the recombination patterns in chicken genome and allowing fine mapping of QTLs (Groenen *et al.*, 2009). Many QTL analysis studies have been conducted on chicken leading to the detection of 3,442 QTLs for 286 different traits which have been curated in the chicken QTL database (<http://www.animalgenome.org/cgi-bin/QTLdb/GG/summary>). A genome-wide physical map of chicken consisting of 2,331 overlapping BAC (Bacterial Artificial Chromosomes) contigs has been constructed providing a powerful platform for research in many areas of chicken genomics such as targeted marker development, fine mapping of QTLs, positional cloning, analysis of genome organization and evolution, comparative genomics, and large-scale genome sequencing (Ren *et al.*, 2003). Over 600,000 chicken Expressed Sequence Tags (ESTs) are now available in the current version of dbEST (release 120701). All these resources have further uplifted the status of chicken as a model species. Furthermore, chicken is ideal for biological research for many other reasons such as the ease of maintaining chicken flocks, their rapid reproduction, large family size, and the availability of generations of pedigree and phenotypic records for maintained and farm populations (Schmutz and Grimwood, 2004).

Genetic and genomic research has been revolutionized with the development of high-throughput genome analysis tools which allow simultaneous analysis of many genomic regions. One of the most important high-throughput technologies is the DNA microarray which has been applied to a diverse range of studies such as for transcriptome analysis, detection and characterization of genetic variants (e.g. Single Nucleotide Polymorphism (SNP), Indels and Copy Number Variants (CNVs)), studying DNA-protein interaction, and

detecting genome methylation. A DNA microarray is defined as a collection of microscopic DNA spots attached on to a solid surface (such as glass, plastic or silicon slides) where each spot contains a minute quantity (picomoles) of a specific DNA sequence called a probe (http://en.wikipedia.org/wiki/DNA_microarray). There are, however, other variants of microarrays where the probes are attached to small microspheres (or beads of 3 to 5 μm in diameter) rather than on to a fixed surface (Tsuchihashi and Dracopoli, 2002). Each bead contains thousands of copies of a specific probe for capturing the target. These beads may remain suspended in assay solution (Luminex platform) or be captured in solid wells (Illumina platform). Unlike in the chip-based arrays where the position of each probe-specific DNA spot is already known, in bead arrays the probe-specific beads are randomly positioned (both when in solution or when deposited on a solid surface) and the identity of each bead is decoded using either different concentrations of fluorescent dyes or by barcoding the probes using some unique sequence (Tsuchihashi and Dracopoli 2002; <http://www.ncbi.nlm.nih.gov/projects/genome/probe/doc/TechBeadArray.shtml>). By containing thousands of probes, a single array can perform massively parallel high resolution analysis of many genetic sites. Although microarrays have been developed on many different platforms (e.g. Affymetrix, Illumina, Luminex, Agilent, Nimblegen etc.) or for different purposes (e.g. gene expression profiling, genotyping, CNV analysis etc.), the underlying basic principles of their functioning are same and these are: (1) hybridization between the target and the probe sequences due to their complementary nature and (2) fluorescent or chemi-luminescent detection of hybridization signal using sophisticated laser instrument for further analysis. Apart from the use of chip-based or bead-based arrays, the microarray platforms also vary on many other aspects such as the type of probes used (cDNA probes, PCR product or oligonucleotide probes), whether the probes were pre-synthesized or in situ synthesized, method of printing/deposition/attachment of probes to glass slides or beads, methods of target DNA preparation, nature of fluorescent dyes, hybridization approach, the nature of scanner used for interpretation of hybridization signal, and the type of algorithms used for analyzing data (Hardiman, 2004; Miller and Tang, 2009; Yauk and Berndt, 2007). Some good reviews on different microarray technologies, their characteristics, advantages and limitations are provided by Tsuchihashi and Dracopoli (2002), Hardiman (2004), Yauk and Berndt (2007), and Miller and Tang (2009).

Since its emergence in 1990s, the microarrays have played a paramount role in genetic and genomic research in human and other model organisms including chicken. However, as the technological advancements have progressed, massively parallel sequencing

technologies have now emerged with the promise to revolutionize, yet again, genetic and genomic research. With the advent of these next generation sequencing (NGS) technologies, many are predicting the end of the microarray era. In this paper, we would like to review the existing microarray resources for chicken, discuss their past and potential future applications, and shed some light on the future of microarray-based research in chicken in the NGS era. We argue that in spite of certain advantages of NGS, the microarrays will continue to exist as major tools in livestock genetic and genomic studies due to their lower price, high throughput nature, established mature technology, and ease of application. We envisage that NGS and microarrays will complement each other rather than substitute.

The rest of this review article is structured around three main sections. The first section, consisting of several subsections, reviews the development and usage of different chicken microarrays viz. gene expression arrays, SNP genotyping arrays, CGH arrays and capture arrays. The second section compares the microarrays and NGS based on their limitations and advantages to discuss the future of array based research and practice. Finally the article ends with a brief section with concluding remarks.

MICROARRAY RESOURCES IN CHICKEN AND THEIR APPLICATIONS

Most of the DNA microarrays that have so far been developed in chicken are either gene expression arrays, genotyping arrays or CGH arrays for characterizing CNVs. Only a recent study has reported the use of capture array technology for enrichment of targeted regions in chicken genome for follow-on sequencing. In the following sections we review the major microarray resources for chicken and their various applications.

Gene Expression Microarrays

Gene expression microarrays are used to study the transcriptome profiles of organisms and have been used for different purposes such as for discovering gene function, elucidating genetic regulation and biological pathways underlying specific physiological conditions, investigating hosts' response to pathogenic infections, disease diagnosis and drug discovery (Murphy, 2002). The basic principle of array-based gene expression analysis is quite simple and is based on comparison of expression patterns between two samples. The entire mRNA content from both the samples are first extracted, purified and then converted to cDNA (Murphy, 2002). The two cDNAs are then labeled with different fluorescent dyes and are hybridized to probes present in the array. The hybridization signals and intensities are then analyzed to interpret differential expression of genes in the two analyzed samples. The probes

for gene expression arrays can either be cDNA or oligonucleotides. Although in the early stages, the cDNA based arrays were more prevalent but due to problems with annotation, clone identity and probe performance associated with these arrays, oligonucleotide (50–70 mer) platforms increasingly became more popular (Woo *et al.*, 2004). Comparisons showed that oligonucleotide arrays offer several advantages over cDNA platforms in terms of specificity, sensitivity, and reproducibility (Hughes *et al.*, 2001).

Over a dozen microarrays (either cDNA or oligonucleotide arrays) have been developed for chicken to study the gene expression profiles (Cogburn *et al.*, 2007). Most of the early stage microarrays were developed in-house and were custom-made for specific studies such as a 3K cDNA lymphocyte array was developed to investigate the gene expression during *myc* oncogene-induced lymphomagenesis in the Bursa of Fabricius (Neiman *et al.*, 2001); a 8K cDNA array specific for pineal gland was developed to study the transcriptional profiles related to the circadian rhythms of melatonin biosynthesis (Bailey *et al.*, 2003); a retina specific cDNA microarray was used to identify and characterize genes expressed in the retinas of chicken embryos to serve as candidate genes involved in the development and function of photoreceptors and other retinal cell types (Hackam *et al.*, 2003); a cDNA array specific to Bursa of Fabricius was used to detect changes in gene expression in this organ during the development of neoplasm due to the presence of a mutant *c-myc* locus (Neiman *et al.*, 2003); a 11K cDNA heart specific array was constructed to study the genes expressed specifically in chicken heart progenitor cells (Afrakhte and Schultheiss, 2004); a 5K cDNA neuroendocrine microarray was developed based on the RNAs expressed in the hypothalamus, anterior pituitary gland and pineal gland of embryonic and young chickens to identify genes that are potentially involved in proliferation and differentiation of cell types in the anterior pituitary gland during embryonic development (Ellestad *et al.*, 2006).

Initially, one of the major setbacks in developing high density gene expression arrays for chicken was the lack of ESTs from different tissues and organs of chicken. In order to boost the functional genomics research in chicken, a number of major projects were undertaken to develop cDNA libraries and sequence ESTs. For instance, in the year 2000, a major functional genomics project for chicken was initiated with the support from USDA-IFAFS to generate a comprehensive catalogue of tissue specific ESTs and develop a number of high density tissue-specific and system-specific gene expression microarrays (Cogburn *et al.*, 2003). Five primary and normalized chicken cDNA libraries from liver, abdominal fat, breast and leg muscle epiphyseal growth plate, neuroendocrinal system (pituitary/hypothalamus/

pineal), and reproductive tract were constructed. Over 30,000 clones from these libraries were sequenced for generating ESTs. One liver specific and two high density multitissue DNA microarrays (representing genes from metabolic/Somatic system and neuroendocrine/reproductive system) were constructed under this USDA-IFAFA project using non-redundant ESTs. The metabolic/somatic Systems microarray was constructed using a set of 11K non-redundant ESTs originally sequenced from cDNA libraries from liver, abdominal fat and skeletal muscle/epiphyseal growth plate. The Neuroendocrine/Reproductive System microarray (8K) was developed using ESTs sequenced from cDNA libraries from the pituitary gland, hypothalamus, pineal gland and reproductive tract including oviduct, ovary and testis. The three arrays developed from this project were used for global gene expression profiling in two populations of broiler chickens divergently selected (1) for extremes in growth rate i.e. a fast growing line (FGL) and a slow growing line (SGL), and (2) for high and low fat contents at a similar growth rate (fat line, FL and lean line, LL). These two system-specific microarrays, however, were later combined to develop a new integrated microarray, called the Del-Mar 14K Chicken Integrated Systems microarray (Cogburn *et al.*, 2004, 2007).

Another large EST sequencing project was initiated by a consortium of the Biotechnology and Biological Sciences Research Council (BBSRC), University of Delaware (UD) and Fred Hutchinson Cancer Research Center (FHCRC) (Boardman *et al.*, 2002; Burnside *et al.*, 2005; Cogburn *et al.*, 2003). The BBSRC project generated in total of 339,314 ESTs from 64 cDNA libraries representing a wide range of embryonic and adult tissues (<http://www.chick.manchester.ac.uk/>; Boardman *et al.*, 2002). The UD Chick EST project had a major focus on the immune system while the FHCRC EST collection was generated from cDNA from chicken Bursal cell lines (Burnside *et al.*, 2005; Caldwell *et al.*, 2005). By combining resources from all these sources, a multitissue cDNA microarray with 13,007 features including 160 control spots was developed. This array provided a broad coverage of the genes expressed in varieties of tissues and also of tissue specific genes. The array is available for academic researchers from genomics@fhcr-c.org (Burnside *et al.*, 2005).

For chicken several microarrays were developed to study gene expression related to immune response. For instance a 14K cDNA Macrophage specific array (Bliss *et al.*, 2005) was constructed to obtain a better insight into the functions of avian macrophages which are critical components of the immune system and plays important roles in both innate and acquired immune responses. This array has been used to examine the transcriptional response of chicken macrophages to Gram-negative bacteria and their cell wall components

and to evaluate the contribution of the Toll-like receptor (TLR) pathway. A well-annotated 5K cDNA immune array was developed by the Roslin Institute for examining host immune response in relation to various avian diseases (Smith *et al.*, 2006). The array was created from libraries developed from a pool of stimulated immune tissues including Bursa, spleen, Peyer's patch, and thymus from chickens which were previously vaccinated against a number of common bacterial, protozoan and viral diseases. The tissues that were chosen for library construction were highly representative of T and B cell populations. The array contained genes that were known to be involved in a wide spectrum of immune function as well as genes previously unknown. This array is available from the ARK Genomics resource centre (<http://www.ark-genomics.org/>). These immune related microarrays provide robust platforms for characterizing host response in chicken against invasion of a variety of pathogens.

Within the last few years, a number of oligo-based microarrays have been developed for whole genome transcriptome profiling in chicken and have been made commercially available. These include a 20K Roslin/ARK CoRe Array V1.0 genome array (Operon Roslin/ARK 2007), a 33K Affymetrix GeneChip microarray (Affymetrix 2007), and a 44K Agilent whole genome transcriptome microarray (Li *et al.*, 2008). The 20K array from Roslin/ARK is a long oligonucleotide (70-mer) array representing 20,673 transcripts. This array is available from Operon (<https://www.operon.com/>) in ready-to-spot 384-well plates or as printed arrays from the University of Arizona (<http://www.grl.steelecenter.arizona.edu/>) and ARK Genomics (<http://www.ark-genomics.org/>) (Cogburn *et al.*, 2007). The Affymetrix 33K GeneChip® array is a short oligonucleotide (25-mer) array with a comprehensive coverage of 32,773 chicken transcripts along with 684 viral transcripts from 17 different avian viruses. Sequence information for developing the array was selected by searching most of the major public databases such as GenBank®, UniGene, and Ensembl (http://www.affymetrix.com/estore/browse/products.jsp?productId=131426#1_1). The 44K Agilent microarray (Li *et al.*, 2008) is, so far, the most comprehensive gene expression platform developed for chicken. The array consists of 42,034 oligonucleotide probes (60-mer), which have been designed from all potential genes based on the chicken reference genome (WASHUC2.1). The probes represent 28 autosomes, two sex chromosomes, the unlocalized chromosomes and mitochondria, plus 1,264 positive control features and 153 negative control features. This array also includes probes designed from 150 chicken microRNA, 43 Marek's disease virus genes and 20 avian influenza virus genes (10 H5N2 and 10 H5N3 genes). These three commercially available arrays have mostly replaced the in-house microarrays as they offer more standardized

platforms for gene expression studies and hence, offer better reproducibility (Cogburn *et al.*, 2007).

The above described microarrays have made invaluable contributions in advancing functional genomics research in chicken by providing insight into metabolic and regulatory pathways associated with different physiological processes and phenotypes, roles of different organ-systems in controlling various biological processes at different phases of life, mechanisms of host-pathogen interaction and the roles of immune system components in fight against pathogenic invasion, gene-environment interaction etc. (Cogburn *et al.*, 2004, 2007). Table 1 provides brief descriptions of some of the selected recent studies on chicken using gene expression microarrays showing their diverse usage. Some good reviews on gene expression studies using microarray techniques in chicken can be found in Cogburn *et al.* (2004, 2007).

Apart from their classical use in elucidating gene expression profiles under various genetic and environmental conditions, the array based transcriptome analyses have been combined with genetic linkage analysis to identify expression QTLs (eQTL) in different organisms, leading to the development of a new integrative field called Genetical Genomics (de Koning *et al.*, 2007). This approach treats the expression levels as quantitative traits and aims to find the genetic variants that influence gene expression (de Koning *et al.*, 2007). The detected eQTL can be either *cis*-acting (that lies close to the gene being controlled) or *trans*-acting (unlinked to the gene being controlled), which provide valuable information regarding the genetic pathways underlying complex trait variation. Recently, a few studies (Blum *et al.*, 2011; Le Bihan-Duval *et al.*, 2011; Le Mignon *et al.*, 2009) on chicken have shown the usefulness of integrating the transcriptome data with QTL analysis. Le Mignon *et al.* (2009) applied three different but complementary approaches to combine the transcript data from a 20K chicken oligonucleotide array to improve the characterization of a chicken QTL for Abdominal Fat (AF) previously detected on distal region of chromosome 5 (GGA5). The first approach dissected the AF phenotype by identifying animal subgroups according to the gene expression profile of 660 transcripts. This led to the identification of a novel QTL in the middle of GGA5 and increased the significance of distal GGA5 QTL, thereby refining its location. The second approach aimed to identify genes which were potentially correlated with the AF trait and regulated by the GGA5 AF QTL region. Five of the 660 genes were considered to be controlled either by the QTL itself or a mutation near to it; one of those being related to lipid metabolism. A QTL analysis with a multiple trait model combining this five gene set and AF was performed which allowed refining the QTL region. The third approach used the transcriptome profiles of these five

genes to predict the paternal Q versus q AF QTL mutation for each recombinant offspring. This approach could reduce the confidence interval of QTL location from 31 cM (with 100 genes) to 7 cM (12 genes). Another example of successful integration of classical QTL analysis and gene expression QTL in chicken was to identify causal gene or QTG underlying a highly significant QTL controlling the variation of breast meat color in a F2 cross between divergent high growth and low growth chicken lines (Le Bihan-Duval *et al.*, 2011).

These results show that the gene-expression arrays are extremely valuable resources for advancing our understanding of chicken genome. Moreover, the ESTs and cDNA sequences that have been generated to create these arrays are also extremely important resources which have facilitated the annotation of the genome improving the status of chicken as a model species (Cogburn *et al.*, 2007).

Genotyping Array

While gene expression microarrays completely changed the dynamics of genomic research in 1990s, the use of genotyping arrays (also called genotyping assays) for high throughput genome-wide analysis came into play much later due to a number of bottlenecks associated with distinguishing allelic change at single-base resolution and multiplexing the reactions over thousands of markers (Syvanen, 2005; Tsuchihashi and Dracopoli, 2002). However, over the past ten years tremendous progress has been made in developing robust platforms for high throughput genotyping arrays and this has revolutionized the way how we can now investigate the genetic architectures of animal and plant genomes. The currently available major genotyping platforms either use allele-specific hybridization approach (Affymetrix) or single-base primer extension method (Illumina). In the first method, the distinction between SNP alleles is achieved using the difference in thermal stability between a perfectly matched and mismatched allele-specific oligonucleotide (ASO) probe and its target sequence (Syvanen, 2005). The thermal stability depends on several factors such as the sequence context of the SNP and the stringency of the reaction conditions including the temperature and ionic strength of assay solution. With this approach, it is however, difficult to achieve optimal hybridization across all the probes. To overcome this problem, SNPs for the array are carefully selected on the basis of their predicted performance in the assay and also several redundant probes are used to interrogate each SNP. Genotype is determined by analyzing differential hybridization of the target DNA over all the redundant probes. The primer extension method on the other hand, does not depend on the sequence context of the SNP and hence require fewer probes per assay (Syvanen, 2005). This

Table 1
Some Selected Recent Studies on Gene Expression Analyses in Chicken Using Microarrays

Reference	Array used	Description of study
(Wang <i>et al.</i> , 2007)	Affymetrix GeneChip Chicken Genome Array	Excessive accumulation of lipids in the adipose tissue is a major problem in the broiler industry. This study analyzed the expression of adipose tissue genes that are involved in pathways and mechanisms leading to adiposity in chickens. Chicken Genome Array was used to investigate differentially expressed genes (DEG) from 7-week-old broilers from lean and fat lines divergently selected over eight generations for high and low abdominal fat weight. Depending on the individual bird, gene expression profiles detected 13,234–16,858 probe sets in adipose tissue. Genes involved in lipid metabolism and immunity such as fatty acid binding protein (FABP), thyroid hormone-responsive protein (Spot14), lipoprotein lipase (LPL), insulin-like growth factor binding protein 7 (IGFBP7) and major histocompatibility complex (MHC), were highly expressed. In contrast, some genes related to lipogenesis, such as leptin receptor, sterol regulatory element binding proteins 1 (SREBP1), apolipoprotein B (ApoB) and insulin-like growth factor 2 (IGF2), were not detected. While most of the previous studies on lipid accumulation focused on liver, this is the first investigation on adipose tissues providing the groundwork for further understanding of the basic genetic control of growth and development of chicken adipose tissues.
(Chiang <i>et al.</i> , 2008)	44K Chicken Agilent Oligo microarray	Contaminated poultry products are the usual sources of salmonella infection and <i>Salmonella enterica</i> serovar Enteritidis (SE) is a major serotype in many countries. Previous studies on SE resistant and SE susceptible broiler lines have shown that heterophils (a type of leukocyte) play important roles in defence against SE infections. This study investigated the DEG profile in heterophils from these two genetically divergent lines following in vitro stimulation with SE with the goal to dissect the interplay between heterophils and SE infection. More DEGs were found between different lines than between infected and non-infected control samples within line. The immune-related genes were more strongly unregulated in resistant line heterophils compared with those in susceptible line, and these genes include several components in the Toll-like receptor (TLR) signaling pathway, and in T-helper cell activation. The study concludes that higher expression of immune-related genes might be more beneficial to enhance host immunity in the resistant line.
(Kano <i>et al.</i> , 2009)	44K Chicken Agilent Oligo microarray	Using transcriptional analysis, this study aimed to understand the protective mechanism provided by Marek's disease (MD) vaccine, produced from antiviral strain of MD virus (MDV). The gene expression analysis was conducted in MDV-infected chickens with and without vaccination at 7 and 21 days post-infection (dpi). The expression data suggested that CD8 α (high) T cell receptor (+) 1 cell population is probably one of the key factors involved in the protective mechanism induced by the viral strain of MD vaccine.
(Bureau <i>et al.</i> , 2009)	21K Operon Roslin/ARK CoRe Array	Components of the Hypothalamic-Pituitary-Adrenal (HPA) axis are known to be important in re-establishing homeostasis following stress. The adaptability response to stress varies considerably among different species, breeds and even individuals within a population as a result of genetic diversity and HPA axis component may play a role in this variability of response to stress. The aims of the present study were to identify genes involved in the regulation of adrenal activity following ACTH (Adrenocorticotrophic hormone) stimulation and to examine differentially expressed adrenal genes in individuals with high and low plasma corticosterone response following ACTH treatment. Gene expression analysis in this study, indicated that ACTH treatment affected the expression of 134 genes. Several genes involved in the adrenal ACTH signaling pathway and steroidogenic enzymes were found to be differentially expressed by ACTH treatment. Quantitative RT-PCRs

TABLE . Continued

Reference	Array used	Description of study
(Heidari <i>et al.</i> , 2010)	Affymetrix GeneChip Chicken Genome Array	<p>were used to validate 18 genes. Only 4 genes revealed to be differentially expressed between higher and lower adrenal responders to ACTH treatment. The results from the present study reveal putative candidate genes and their role in regulation of adrenal functions.</p> <p>Marek's disease (MD) is a lympho-proliferative disease of chickens induced by alpha-herpesvirus called Marek's disease virus (MDV). This study used a global host gene expression analysis in the splenocytes (spleen lymphocytes) of MDV-infected chickens to gain insight into the molecular mechanisms associated with MDV infection and the chicken gene expression pattern in response to the infection. Two-week-old MD-susceptible chickens were challenged with an oncogenic strain of MDV, and spleen samples were collected 5 and 15 days post-infection (dpi) for RNA isolation. Array data showed a significant differential expression of immune response genes between the two phases of MDV infection. At 5dpi, the expression levels of more than 22 immune-related genes were downregulated and at least 58 genes were upregulated, compared with control birds of same age. In comparison, at 15 dpi (latency infection), 67 immune-related genes were downregulated and only 6 genes were upregulated. Cytokines, chemokines, MHC molecules and related receptors, and adhesion molecules were among the many MDV-induced downregulated genes that are critical for an effective antiviral immune response. In addition, several apoptosis-associated genes were decreased in expression during latent infection, suggesting an MDV-induced blocking of initiation or progression of programmed cell death processes.</p>
(Smith <i>et al.</i> , 2011)	Affymetrix GeneChip Chicken Genome Array	<p>This study used comprehensive gene expression analysis to study the host response to MDV infection in both MD resistant and susceptible chicken lines to identify genes and pathways involved in susceptibility to the disease. The study suggested a novel pathogenicity mechanism involving the downregulation of genes containing H1C1 transcription factor binding sites as early as 4 days post-infection. Since H1C1 drives anti-tumor mechanisms; so its downregulation suggests that MDV infection switches off genes involved in antitumor regulation several days before the expression of the MDV oncogene. The comparison of the expression data to previous QTL data identified several candidate genes for involvement in resistance to MD. One of these genes, IRG1, was confirmed by SNP analysis to be involved in susceptibility. The gene expression analysis suggests that this gene probably has a role in apoptosis.</p>
(Kong <i>et al.</i> , 2011; Bottje <i>et al.</i> , 2012;)	44K Chicken Agilent Oligo microarray	<p>These two papers studied DEG in breast muscle of individuals phenotyped for high or low feed efficiency (FE) within a single male broiler line. Among the 782 DEGs, the first study only looked into the top 10 unregulated and 10 downregulated genes and found that the upregulated genes in the high-FE group were generally associated with anabolic processes. In contrast, most of the downregulated genes in the high-FE group were associated with muscle fiber development, muscle function, and cytoskeletal organization, while 3 genes were associated with self-recognition or stress-response. The 2nd paper, investigated 27 focus genes that were selected using the IPA (Ingenuity Pathway Analysis) software with the goal to identify genes that could play critical roles in FE based on frequency of appearance in key biochemical (canonical) pathways. Focus genes that were upregulated in the high FE phenotype were associated with important signal transduction pathways (Jnk, G-coupled, and retinoic acid) or in sensing cell energy status and stimulating energy production that would likely enhance growth and development of muscle tissue. In contrast, focus genes that were upregulated in the low FE muscle phenotype were associated with cytoskeletal architecture (e.g., actin-myosin filaments), fatty acid oxidation,</p>

TABLE . Continued

Reference	Array used	Description of study
(Habig <i>et al.</i> , 2012)	Affymetrix GeneChip Chicken Genome Array	growth factors, or ones that would likely be induced in response to oxidative stress. The results of these studies provide valuable insight into the gene expression and the cellular basis of FE in broilers. This study investigated the DEG in chickens from a white egg layer line (Lohmann Selected Leghorn, LSL) and a brown egg layer line (Lohmann Brown, LB), when the hens of these lines were kept in the newly developed small group housing system (Eurovent German) with two different group sizes. Even though these two chicken lines have similar egg laying performance, they differ considerably in many other phenotypic and behavioral traits. Whole genome RNA expression profiling was used to study the differences in gene expression profile under this new housing system. Differential expression was observed for 6,276 array probes but a twofold or greater change in gene expression was identified for 151 probes. In LSL, 72 of the 151 probes were upregulated and 79 downregulated. Majority of the upregulated genes were found to be related to immune system processes and membrane organization. Majority of the downregulated genes were related to phosphorus metabolic processes and signaling pathways.
(Li <i>et al.</i> , 2012)	Custom 9K cDNA array from breast and liver tissue	Roxarsone, an animal growth additive, significantly improves the growth of broiler chicken but its application not only contaminates the animal product but also the environment. This study was conducted to understand the response of genes to roxarsone as this knowledge may facilitate the discovery of new, safer substitute. Studying 8,935 genes in chicken breast muscle using a cDNA microarray showed that 30 genes have consistently up- or downregulated throughout the medication periods. Of these 30 genes, 13 were well documented of which 11 genes were related to immunity while the remaining two genes were related to energy metabolism. This finding suggested that the roxarsone supplement probably promote growth by improving immunity of chicken through regulation of immunity related genes.

method uses two bead types (each with allele-specific probes) per SNP assay whose probe sequences differ only at the 3' terminal base. The probes hybridize with the target DNA immediately adjacent to the SNP site and allele discrimination is achieved by polymerase extension step of a single base. Primer extension and fluorescence labeling occurs only when the probes have hybridized to a perfectly matched allelic targets (Gunderson *et al.*, 2005).

The most important use of genotyping arrays have been in mapping of human disease susceptibility loci as can be seen from the catalogue of published genome-wide-association studies (GWAS) (Hindorff *et al.*, 2009). A detailed human haplotype map has been created using over a million SNP to facilitate the GWAS (The International HapMap Consortium, 2005). This HapMap is not only a valuable resource for designing GWAS but also for detecting recombination hotspots across the human genome, getting insight into the effects of natural selection during human evolution and shedding light on the presence of structural variations (SVs). Dense genotyping arrays have now been developed for many other organisms including farm animals and have been

used for a variety of genetic and population genetic studies including GWAS, detecting linkage disequilibrium (LD) structure of the genome, discovering population structure, tracking the footprints of natural and domestication selection, and in livestock animals as an aid to breeding and selection.

For chicken, traditionally the quantitative and population genetic studies were conducted using a handful of markers, generally microsatellites. Developing genotyping arrays using hundreds and thousands of SNP markers was not feasible earlier due to an absence of a large collection of genome-wide SNPs. This scenario, however, changed after the release of the first draft of chicken genome sequence (WASHUC1.0) in the year 2004 (International Chicken Genome Sequencing Consortium, 2004). This sequence assembly allowed creating a genetic variation map for chicken with 2.8 M SNPs and Indels by comparing the sequence of three domestic chicken breeds—a broiler, a layer, and a Chinese Silkie—with their wild ancestor, the red jungle fowl (International Chicken Polymorphism Map Consortium, 2004). Although a single individuals from each line were compared for detecting the variants, subsequent

experiments suggested that at least 90% of these were true SNPs, and at least 70% are common SNPs that segregated in many domestic breeds (International Chicken Polymorphism Map Consortium, 2004). The very first chicken SNP array was developed by a consortium under a USDA funded project by selecting 3,072 SNPs that were evenly distributed throughout the genome from the collection of 2.8 M chicken SNPs (Muir *et al.*, 2008b). To select the SNPs, the genome was divided into 3,072 bins taking into account the recombination rate per chromosome. Within each bin, three SNPs were selected. Preference was given to high confidence SNPs within genes, especially those predicted to be exonic non-synonymous SNPs but does not disrupt the protein deleteriously (i.e. predicted to be “tolerated” in SIFT analysis). The selected SNPs were evaluated for Illumina assay suitability and the best SNP within each bin was selected for developing the array. The 3,072 finally selected SNPs were validated by genotyping 1,440 birds from 24 elite commercial lines (60 birds/line). A total of 2,733 SNPs (89%) converted successfully with a reproducibility rate as high as 99.996%. The SNPs used for this array can be downloaded from <http://poultry.mph.msu.edu/resources/resources.htm#SNPs>. This 3K array has been used in several genetic studies such as for updating the chicken genetic map which in turn facilitated the creation of the second assembly (WASHUC2.1) of the chicken reference genome (Muir *et al.*, 2008b); for genome wide association analysis for detecting QTLs contributing to fatness related heterosis observed in F2 populations created by crossing two divergent chicken stocks (Abasht and Lamont, 2007); and for investigating extant genetic diversity in commercial chicken breeds compared with ancestral breeds (Muir *et al.*, 2008a).

Following the development of this first genotyping array, many other studies have reported using low to medium density (from >1K to 60K) SNP arrays for various purposes; many of these were extensions of the original 3K panel with more SNPs added. For instance, a 6K panel was developed by adding another 3,000 SNPs to the original 3K panel to obtain a greater density to investigate the extent and consistency of LD in related chicken lines (Andreescu *et al.*, 2007). This study analyzed 959 SNPs from chromosome 1 and 398 SNPs from chromosome 4 on 179 to 244 individuals from each of nine commercial broiler breeding lines and the results showed that LD (r^2) extended over shorter distances than reported previously in other commercial livestock populations. Moreover, the LD at short distance (within 1 cM) tended to be consistent across related populations, thereby, suggesting that LD based information (such as marker-QTL association) from one line would be usable in related lines.

An Illumina Golden Gate/Sentrix array was developed with 889 SNPs selected from dbSNP 125 from

several chromosomal regions, to study the LD and haplotype diversity in macro- and microchromosomes in chicken (Megens *et al.*, 2009). The SNPs were selected from several ~1 cM regions on macrochromosomes, GGA1 and GGA2, and 1.5 to 2 cM regions on microchromosomes, GGA26 and GGA27. The analysis of these SNPs on 371 birds from eight chicken populations (including both commercial and traditional breeds) provided a number of valuable insights into the genetic architectures of chicken populations. First, it confirmed that at a similar physical distance, LD, haplotype homozygosity, haploblock structure, and haplotype sharing were all lower for the microchromosomes compared with those in macrochromosomes and that this pattern was consistent across breeds. Second, there were differences in LD, haplotype variation, and haplotype sharing between populations and such difference was shaped by the known demographic history of the populations. Like other previous studies, this also showed that the extent of LD was greater in commercial layer lines compared with broilers. Third, even though the extent of LD overall was very low, some haploblock structures (typically expanding <10 kb) were observed, particularly in the macrochromosomes. The authors concluded that due to the limited haploblock structure and LD, the future whole-genome marker arrays will require >100K SNPs to exploit all the haplotype information. Besides, the effective interpretation and transfer of genetic parameters will require taking into account the size of the chromosomes.

Groenen *et al.* (2009) used several Illumina Golden Gate assays to genotype 12,945 SNPs to create a consensus linkage map for chicken with greater resolution than was previously available. They used three marker panels for this purpose: the original 3K panel, a second panel consisting of 657 SNPs which were not assigned to a chromosomal location on Build WASHUC1.0, and a third panel consisting of 9,216 SNPs selected to be evenly spaced throughout the chicken genome, also taking into account the contigs that were not yet assigned to any chromosomes on Build WASHUC1.0. By applying sex-specific analysis and comparison of the linkage maps in three individual populations, this study observed prominent heterogeneity in recombination rates between populations but no significant difference between sexes. A further high resolution linkage map was created in another study by using a 18K Illumina Infinium iSelect Beadchip and by genotyping 1,619 birds from two chicken populations: a purebred broiler line and a broiler X broiler cross (Elferink *et al.*, 2010). This study observed some regional differences in recombination hotspots between the two mapping populations in several chromosomes, even though the overall recombination pattern was quite similar in these populations. A sex-specific analysis revealed that these

regional differences were originated mainly due to female-specific recombination hotspots in the broiler × broiler cross.

One major area of application of high density genotyping array in farm species is the Genomic Selection (GS), which is the most recent advancement in the marker-assisted selection strategy. GS uses genotypes from dense marker panels for estimating total genomic breeding value (GEBV) (Meuwissen *et al.*, 2001) of animals for selection traits. Rather than directly searching for functional units associated with each trait—which for quantitative traits may be many in numbers and are difficult to identify exhaustively—GS incorporates in the GEBV calculation the effect of all the markers on the target trait, considering that all possible QTLs have been tagged by at least one marker. GS promises to accelerate genetic gains by capturing all or most of the genetic variance present in the genome and by facilitating selection at an early stage leading to shortening of the generation interval (Meuwissen *et al.*, 2001). GS is being applied extensively for genetic improvement of dairy cattle and two major high density arrays (>500K) are now available for cattle (Rincon *et al.*, 2011). For chicken, several studies have reported the use of proprietary medium density arrays for GS in commercial broiler and layer breeds (Avendaño *et al.*, 2010; Dekkers JCM, 2010; Preisinger, 2012; Wolc *et al.*, 2011). For instance, the EW group of poultry breeding companies (including Aviagen Ltd., Hy-Line International and Lohmann LTZ) developed a proprietary 42K Illumina iSelecta BeadChip for application to GS (Avendaño *et al.*, 2010). They optimized the array to extensively capture the genetic variance associated with a range of economically important traits and the LD structure within and between broiler and layer populations. In 2011, a 60K well-characterized chicken genotyping array was developed under a USDA funded project and in collaboration with Wageningen University and two poultry breeding companies, Cobb-Vantress, USA and Hendrix-Genetics, The Netherlands (Groenen *et al.*, 2011). This Illumina SNP BeadChip array consisted of 57,636 SNPs of which 54,293 could be validated through genotyping and were shown to be segregating in chicken populations. This project applied the Illumina NGS approach on reduced representation libraries to detect novel segregating SNPs not present in public database to achieve a better coverage of the genome. Moreover, they identified and included SNPs from contigs that were not covered by the existing chicken genome assembly (WASHUC2.1). Although this array has been used in a number of genetic analyses of diverse nature such as the analysis of LD to understand the demographic history in chicken breeds (Qanbari *et al.*, 2010), detection of selection signatures in a wide varieties of chicken breeds (Elferink *et al.*, 2012) and genome-wide association study of body weight (Gu

et al., 2011), the array has limited availability as it has not been commercially released for wider use.

To fill in the vacuum of a lack of commercially available genotyping array for chicken, recently a high density Affymetrix Axiom array has been developed under a BBSRC/DEFRA LINK grant project with collaboration among the Roslin Institute, several poultry breeding companies (Aviagen Ltd., Hyline International), Affymetrix Ltd. and the German Synbreed consortium (Kranis *et al.*, 2013). The project was initiated with the goal to develop a 600K SNP array in such a way that it provides a comprehensive coverage of the chicken genome and can be used for whole genome analyses such as for GS and GWAS in a variety of chicken breeds. In order to achieve this, the project undertook a massive effort of re-sequencing many birds from a variety of commercial and experimental chicken breeds to detect a large catalogue of segregating SNPs. In total, 243 chickens from 24 lines (4 broilers, 12 layers, and 9 inbred experimental lines from Institute of Animal Health) were re-sequenced using Illumina NGS technology in combination with sequencing of pooled DNA. By analyzing sequence data from individual lines, over 78M of segregating SNPs were detected with reasonable confidence. A very stringent set of criteria was then applied to filter the SNPs to minimize false positives and to retain only the very high quality markers which provide a uniform coverage of the genome in terms of genetic map distance. Finally, after several steps of filtration ~1.8M SNPs were selected for validation through genotyping before choosing the final panel of 600K SNPs. The final panel consisted of 580,954 SNPs including at least 21,534 coding variants, which had been validated on a large number of chicken populations from a variety of commercial and traditional breeds (3 broiler lines, 5 brown egg layer lines, 3 white egg layers lines, and 13 traditional breeds). Since layers, in general, show a greater extent of LD compared to broilers (Aerts *et al.*, 2007; Andreescu *et al.*, 2007; Megens *et al.*, 2009; Qanbari *et al.*, 2010), this array incorporated a balanced ratio of broiler and layer SNPs (3:2), to provide similar power to capture most of the genetic variance in these two breeds. While the array consists of mostly common SNPs to increase its utility on a wide variety of chickens, it also contains a small proportion of low frequency alleles from different lines. Based on the validation result, this final array contains over 100K to 450K segregating SNPs per line of layer and broiler chickens tested. In a previous study, Megens *et al.* (2009) has suggested that any future assays will require at least 100K SNPs to capture all the genetic variance within a chicken population. The newly developed array, therefore, is expected to be highly effective for whole genome analyses of a wide range of chicken populations. All the SNPs present in this array has been submitted to NCBI dbSNP and can be downloaded from the website

http://www.ncbi.nlm.nih.gov/projects/SNP/snp_view_batch.cgi?sbid=1057286]. This genotyping array can be bought directly from Affymetrix Ltd. or ordered through ARK-Genomics.

Although development of this HD genotyping array has opened up huge opportunities for large scale genome wide analyses, using it for routine applications like GS will still be very expensive. Particularly, the large number of selection candidates that poultry breeding companies need to screen each year (e.g. between 100,000 and 150,000 individuals in a typical broiler elite line), can make the application of HD arrays for GS prohibitively expensive (Avendaño *et al.*, 2010), whereas the use of low density (LoD) panels would be much more affordable. “Genotyping imputation” can surmount this problem by reducing the cost of genotyping while at the same time exploiting the advantage of high density marker genotypes. This approach performs *in silico* prediction of genotypes of study individuals for large number of markers even though actual genotyping has been performed only for a subset of markers (Marchini and Howie, 2010). To apply imputation, a reference population needs to be genotyped for the entire panel of HD array. Imputation then attempts to identify shared haplotype between the study individuals and the reference population and use this sharing to predict the missing alleles in the first group. Genotype imputation, therefore, can considerably improve the power of the study by incorporating many more markers and many more individuals than are possible to be actually genotyped. In the context of GS, the loss-of-accuracy from genotype imputation was investigated through simulation studies (Habier *et al.*, 2009) where ancestral populations were genotyped with HD panel and the descendants including the selection candidate were genotyped using LoD panel with equally spaced markers. The haplotype blocks within the families were used to track the co-segregation of HD and LoD panel SNPs and to predict the genotype of the missing SNPs in descendants. This study showed that losses were small in the first generation of imputation (<5% at an LoD-SNP density of 10 cM) but this increased over generations. However, if the parents in each generation were genotyped with HD panel, then the increased loss-of-accuracy could be prevented. This study, therefore, confirmed that genotype imputation using LoD marker panel can be used for predicting GEBV with only a limited loss of accuracy. Aviagen Ltd is currently applying this method in GS of their broiler stock (Avendaño *et al.*, 2010).

In conclusion, we can predict that this 600K genotyping array will be a highly valuable resource for genetic analyses in chicken by allowing much better resolution compared with the previously used low or medium density panels. In fact, many of the studies like LD structure analysis in different breeds of chicken or in different chromosomes can be repeated to improve the

resolution. The panel can be used to create a haplotype map for chicken which will be a valuable resource with considerable implications in further characterization of chicken genome and in genetic studies like GWAS and GS. Even though only a few million SNPs for chicken were available in public database just a few years ago, two recent studies have reported detection of a large catalogue of variants. One study detected over 7M SNPs by pooled sequencing of individuals from eight domesticated chicken populations and one Red Jungle Fowl population (Rubin *et al.*, 2010). The other major effort of SNP detection was undertaken through the project which developed the 600K chicken genotyping array as described earlier (Kranis *et al.*, 2013). The project team is currently characterizing about 19M good quality SNPs and this would soon be submitted to NCBI dbSNP (personal communication). This large catalogue of SNPs provides a highly valuable resource as their annotation in terms of their genomic positions (exonic, intronic, intergenic, downstream or upstream etc.) and predicted effect (splicing, synonymous, nonsynonymous, stop gain/loss etc.) will allow further characterization of the chicken genome. Furthermore, this SNP resource opens up the opportunity to develop customized arrays for specific purposes such as developing arrays only with coding SNPs or validated rare variants.

Microarrays for Studying CNVs

CNVs—gains, or loss of large genomic segments resulting in abnormal copy numbers of these regions—are now known to be a prevalent type of genetic variant and have been implicated with several diseases and complex traits in human and other animals such as in HIV/AIDs susceptibility (Gonzalez *et al.*, 2005), autoimmune disease (Fanciulli *et al.*, 2007; McKinney *et al.*, 2008), asthma (Brasch-Andersen *et al.*, 2004), Crohn’s disease (McCarroll *et al.*, 2008), Osteoporosis (Yang *et al.*, 2008), etc. In chicken, a CNV in the non-coding sequence in SOX5 gene has been shown to be associated with pea-comb which is an adaptive trait for chicken in cold climates that reduces heat loss and makes the chicken less susceptible to frost lesions (Wright *et al.*, 2009). Another study (Elferink *et al.*, 2008) shows that partial duplication of two genes: PRLR and SPEF2 are associated with late feathering in chicken. These studies indicate that CNV analysis can provide valuable insight into underlying genetic mechanisms of many traits. While traditionally CNVs and other chromosomal abnormalities have been detected using techniques like fluorescence *in situ* hybridization (FISH) and comparative genomic hybridization (CGH), array based technologies have emerged as the most robust methods for genome wide search of CNVs with higher resolution and speed (Carson *et al.*, 2006). Two array based methods for detecting CNVs include the

use of array-CGH and SNP-genotyping arrays (Pinto *et al.*, 2011).

Array-CGH or simply aCGH is the microarray based CGH technique for detecting CNVs. Like CGH method, aCGH uses two genomes—a test and a reference, which are labeled using different fluorophores (Alkan *et al.*, 2011; Theisen, 2008). But unlike the CGH technique, which uses metaphase chromosomes to hybridize the labeled test and control DNAs, the aCGH technique uses probes immobilized on solid surface as the targets for analysis. The probes can vary in size from oligonucleotides (25–85 base pairs) to genomic clones such as bacterial artificial chromosomes (80,000–200,000 base pairs). The ratio of signal intensity (converted to \log_2 ratio) between the test and reference is then used to assess the copy number changes at specific genomic locations. The aCGH technique has been applied to create comprehensive maps of human CNVs (Iafrate *et al.*, 2004; Redon *et al.*, 2006; Sebat *et al.*, 2004; Wong *et al.*, 2007). In recent years, scientist have applied the aCGH or SNP genotyping arrays to investigate the CNV profiles in several non-human organisms, including farm animals such as cattle (Hou *et al.*, 2011; Liu *et al.*, 2010), pigs (Fadista *et al.*, 2008; Ramayo-Caldas *et al.*, 2010; Wang *et al.*, 2012), and sheep (Fontanesi *et al.*, 2011). For chicken, several aCGH have been developed for CNV detection such as 242K and 400K chicken CGH arrays from Agilent, and 385K CGH whole genome tiling array from Nimblegen (Skinner *et al.*, 2009; Volker *et al.*, 2010; Wang *et al.*, 2010, 2012).

The 385K NimbleGen array CGH was used to create the first genome-wide map of CNVs for chicken (Wang *et al.*, 2010). Four Cornish Rock broiler lines, four Leghorn and two Rhode Island Red layer lines were analyzed with the array and a male broiler DNA was used as a reference for all hybridization. Ninety six high-confidence CNVs were identified by analyzing these three lines of chicken and 26 of these CNVs were detected in two or more animals. Whereas most small sized CNVs were found to be present in non-coding regions, the larger ones were found within genes, suggesting that these CNVs can have important association with various traits. This study gave the initial understanding of the CNV profile in chicken genome and the number of CNVs detected per individuals was found to be quite similar to that of several mammalian species. The NimbleGen CGH array has also been used for comparative analysis of Turkey, Duck and Zebra Finch genomes for detecting CNVs (Griffin *et al.*, 2008; Skinner *et al.*, 2009; Volker *et al.*, 2010) demonstrating that such arrays can be useful resources for characterization of other avian species.

Recently, an Agilent 400K whole genome array CGH has been developed for chicken using custom designed probes to create a detailed CNV map for local Chinese breeds and commercial lines (Wang *et al.*,

2012). The array contained a total of 420,288 probes of 60-mer length covering 29 autosomes, two sex chromosomes and 25 random chromosomal fragments. Most of these probes (>98%) were selected from Agilent's High Density Probe Database containing over 4 million validated chicken CGH probes. The rest of the probes included Agilent's positive and negative controls and 450 probes from chrE22C19W28_E50C23, chrE22C19W28_E50C23_random, chrE64 and chrE64_random that were not present in Agilent's CGH probe database. The array provided a comprehensive coverage of exonic, intronic and intergenic regions of the chicken genome with a mean probe spacing of 2,671 bases in WASHUC2.1 chicken build. Using this array, Wang *et al.* (2012) have identified 130 CNV regions with mean length of 25.7 kb, of which 104 segments were reported for the first time. About 54% of these novel CNVs were detected from non-coding regions, 62.5% were gain in copy number while the rest (38.5%) represented copy number loss. The study also detected four regions which have most probably arisen through selection during domestication process of chicken.

Apart from the CGH arrays, SNP genotyping arrays have also proved to be valuable tools for detection of CNVs. Unlike the CGH arrays that compares the hybridization signals from two DNA sources through competitive hybridization, genotyping arrays depends on signal from single source hybridization which subsequently is compared with a set of reference values from control individuals (Carson *et al.*, 2006). SNP microarrays suffer from certain disadvantages but at the same time offer certain advantages compared with CGH arrays. One major disadvantage is that the SNP selection process for designing the array tends to screen out SNPs in CNV regions particularly because they do not convert very well during genotyping. As a result most of the early genotyping arrays showed poor coverage of the CNV regions, although the current arrays perform better (Alkan *et al.*, 2011). Another critical disadvantage is that they tend to provide lower signal-to-noise ratio than do the CGH arrays resulting in greater false discovery rate (FDR) (Alkan *et al.*, 2011). One key advantage of SNP arrays, however, is that they offer the opportunity to calculate a new metric called B allele frequency (BAF) by using SNP allele-specific probes, which can help increase CNV sensitivity, distinguish alleles and identify regions of uniparental disomy (Alkan *et al.*, 2011).

Although HD genotyping arrays have been used for detection of CNV in human and other animals (Hou *et al.*, 2011; Redon *et al.*, 2006; Rincon *et al.*, 2011; Wang *et al.*, 2012), for chicken no studies have so far been undertaken due to lack of such HD arrays. The development of 600K Affymetrix Axiom array now opens up new opportunities for genome-wide screening of CNVs.

Since both CGH arrays and HD SNP arrays now exist for chicken, this offers further opportunity to combine these two platforms for detection of CNVs with greater confidence.

Genomic DNA Capture Arrays

One of the new advancements in the microarray technologies is the “capture arrays” which allows enrichment of specific target regions of the genome for high throughput sequencing. Since whole genome sequencing is still an expensive endeavor, focusing the sequencing effort to the regions of interest is an attractive and cost saving option. Since only the targeted regions are sequenced, the depth of coverage can be increased to achieve a greater resolution of the analyses. While traditionally the enrichment of the candidate region has been achieved using PCR, this approach becomes laborious and time consuming if the number of regions to be enriched is many; for instance, if the goal is to search for many candidate genes for mutations to investigate their association with a trait (Almomani *et al.*, 2011). Array based enrichment can provide a cost-effective and rapid solutions to these issues by eliminating the necessity of performing thousands of PCR reactions, instead allowing for parallel enrichment of target regions in a single experiment (<http://www.nimblegen.com/products/seqcap/index.html>). Capture array, for instance, have been used for targeting and capturing all the protein coding regions in the genome (exome) (Shendure, 2011) or for selected enrichments of candidate genes and regions associated with specific disease or trait (Almomani *et al.*, 2011). This technology is expected to be a major driving force in the next generation of genetic and genomic research in human and non-human species.

Although a relatively new technology, scientists working on the chicken model have already started reaping the advantage of this revolutionary approach. Recently, scientist at the University of California Davis (UCD) has used the targeted capture array technology (Agilent's SureSelect Target Enrichment System) to sequence candidate regions associated with three developmental abnormalities in chicken—*coloboma*, *diplopodia-1* and *wingless-2*—to investigate the underlying causative elements (Robb and Delany, 2012). These three defects are caused by single gene recessive mutations resulting in craniofacial, limb, skeletal, muscular and/or integumental abnormalities, having homologies with developmental defects in human. Congenic inbred lines for these abnormalities were created to facilitate discovery of the specific genetic elements causing each defect. The chromosomal locations and causative regions associated with these defects were previously mapped using SNP genotyping arrays to regions in chromosomes 1, 12, and Z. Due to the congenic nature of

the lines, each targeted region could be compared with the other two congenic partners, thereby providing internal control/reference, all analyzed on a single array. Upon sequencing the enriched captured regions, about 76% of the ~73 million sequence reads were found to be specific to the targeted regions with an average coverage of 132-fold. Analysis of these three targeted regions identified line-specific SNPs, short indels and putative chromosomal rearrangements—all of which are candidate causal variants for these three traits. This is the first report of targeted capture array technology in chicken or in any other avian species.

Although potentially a very powerful technology, exploiting the full potential of DNA capture arrays in genetic/genomic research on chicken would require a completed genome assembly, better annotation of the genome, mapping of transcribed regions, and greater knowledge of the candidate regions to be studied.

FUTURE OF MICROARRAYS IN THE NGS ERA IN CONTEXT OF LIVESTOCK RESEARCH

Even though microarray technologies have revolutionized the genetic and genomic research by allowing high throughput analyses, they inherently suffer from certain limitations. One major limitation of microarrays is that they require *a priori* knowledge of the genome and sequences of genomic regions to be investigated (Hurd and Nelson, 2009; Shendure, 2008; Wang *et al.*, 2009). This affects effective designing of arrays for organisms with incomplete genome sequence or poorly annotated genome (Hurd and Nelson, 2009). Another major drawback of microarrays is that they can investigate only specific regions of the genome which are targeted by probes (Roh *et al.*, 2010), thereby missing any critical regions that are not covered. A major implication of this has been felt in many array-based GWAS analyses which scanned genomes to detect the causative elements underlying various phenotypic or disease trait. Although, most of these studies found multiple SNPs significantly associated with a trait, most of these SNPs explained only a tiny proportion of the genetic variance—leaving most of the heritability unexplained (Maher, 2008). Among many plausible reasons for this missing heritability, failures to account for rare mutations and structural variants (SVs) have been considered as two potential causes (Maher, 2008). In the absence of well characterized catalogues of rare variants in most species, array-based genotyping, in general, have targeted only common SNPs having minor allele frequency >5%. Current microarray platforms also fall short in detecting certain forms of variants such as translocation and inversion types of SVs. Even though CGH arrays or SNP arrays can detect CNVs, resolving breakpoint has often proven difficult (Alkan *et al.*, 2011). Cross-hybridization between similar sequences is another major

limitation of microarrays restricting their focus only to the non-repetitive regions of the genome (Shendure, 2008; Hurd and Nelson, 2009; Wang *et al.*, 2009). This obstructs the analysis of related genes and alternatively spliced transcripts by the gene expression arrays, and the detection of CNVs in repeat-rich and duplicated regions even though CNVs tend to accumulate in such regions in greater frequency (Alkan *et al.*, 2011). Microarrays have suffered from a number of quality issues as well. Gene expression arrays, for instance, have often suffered from poor data quality due to low signal-to-noise ratio, problems with reproducibility of the results, poor sensitivity to rare transcripts and difficulty in detecting alternatively spliced transcripts (Shendure, 2008; Wang *et al.*, 2009). Similarly, CGH arrays also suffer from difficulty in comparing results between different array platforms, poor sensitivity in detecting single copy gains and smaller CNVs (Alkan *et al.*, 2011).

NGS approaches, on the contrary, offers to circumvent most of these limitations. No prior knowledge of the sequence is required and hence organisms with no reference genome or incomplete genome assembly can be studied (Shendure, 2008; Hurd and Nelson, 2009). Most importantly, sequencing offers the potential to uncover novel elements in the genome, for instance, it provides the opportunity to detect all types of variants including SVs and identify alternatively spliced transcripts (Alkan *et al.*, 2011; Hurd and Nelson, 2009; Roh *et al.*, 2010). A number different sequencing approaches have been used for detection of SVs, including the read-pair and split-read approaches that can detect inversions and translocations (Alkan *et al.*, 2011). Sequencing also offers the opportunity to survey paralogous sequences and transcribed repeat elements (Cloonan *et al.*, 2008; Shendure, 2008). While the gene expression microarrays only measure the relative quantity of the transcripts, RNA sequencing (RNA-Seq) can provide an absolute measurement of the transcripts (Mortazavi *et al.*, 2008; Wang *et al.*, 2009). Two papers published in *Nature Methods* in 2008 demonstrated the power of RNA-Seq in detecting and characterizing, with very high resolution, transcripts from both known and unknown genes, alternative splicing and repeat elements (Cloonan *et al.*, 2008; Mortazavi *et al.*, 2008). With all these potential advantages and their rapidly falling prices, massively parallel sequencing technologies have stirred a great deal of debate in the scientific community regarding the future of microarrays in genomic and genetic studies. While many are anticipating “a beginning to the end of microarrays” (Ledford, 2008; Shendure, 2008), others think that sequencing technologies have not yet reached the level to replace microarray; instead they envisage that these two technologies will continue to thrive in future as complementary to each other rather than as competitors (Roh *et al.*, 2010). Here we discuss some of the debates and

reasons why we think the microarrays are going to stay as important genome analysis tools in the foreseeable future particularly in the fields of livestock genetics and genomics.

While NGS approaches promise some major advantages, these are still new technologies having come into limelight only since 2004 and hence would still require substantial developments in terms of improving the technology, quality, and bioinformatics pipelines. Each of the three major NGS platforms—Roche 454 pyrosequencing, Solexa/Illumina Genome Analyser, Applied Biosystems SOLiD—has its own limitations and quality issues. For instances, the Solexa/Illumina platform generates short reads (36–150 bp) making the alignment or de novo assembly of the reads difficult (Hurd and Nelson, 2009). On the other hand, Roche 454 pyrosequencing platform produces longer reads (up to 500 bases) but suffers from inaccuracies in calling homopolymeric stretches of sequences (i.e. AAAAA, CCCCC) due to the system’s chemistry (Hurd and Nelson, 2009). Moreover, all NGS technologies inherently suffer from high error rates compared with Sanger sequencing due to a multiple reasons including base calling, alignment error and low depth of coverage at which the sequencing is performed in most projects (Nielsen *et al.*, 2011). A comparison of Sanger sequencing with NGS platforms showed that although NGS can correctly identify >95% of variant alleles, the average coverage required to achieve this performance is greater than the targeted levels of most of the current studies (Harismendy *et al.*, 2009). Increasing the depth of coverage can reduce the error rate, but this increases the cost drastically. Besides, the current NGS technologies produce highly variable sequence coverage for different regions (Harismendy *et al.*, 2009). Even though the potential in detecting alternative splicing is considered a major advantage of the RNA-Seq approach, this actually becomes quite challenging with short read lengths (Perez-Enciso and Ferretti, 2010). The same issue makes the detection of SVs difficult or resolving their breakpoints challenging particularly when the variant is present in duplicated regions (Alkan *et al.*, 2011).

Apart from the quality issues, cost is still a major issue for NGS. Even though the cost of generating sequencing data is dropping significantly, it is still much more expensive compared to microarrays (Ledford, 2008). Besides, other costs associated with sequencing needs to be taken into account including the costs for sample preparation, bioinformatics analyses, and computing infrastructures for handling and storing large dataset (Perez-Enciso and Ferretti, 2010). Bioinformatics and computational requirements are in fact major obstacles to the fast and widespread adoption of NGS based approaches as most labs lack these resources and skills and establishing these requires substantial investment (Alkan *et al.*, 2011). Bioinformatics pipelines for

sequence analyses are still evolving and have not yet reached maturity as these are confronted with many issues such as optimizing the parameters for sequence alignment, filtration of sequence for quality, de novo assembly of short reads, developing algorithms for detecting short indels etc. In contrast, the bioinformatics pipelines for analyzing array-data are much more mature and established, rapid and relatively easy to use (Perez-Enciso and Ferretti, 2010). The increase in cost associated with sequencing at greater depth of coverage is another issue especially for projects targeting to detect rare variants, or less abundant transcripts or SVs using NGS data. At the current sequencing cost, this would appear prohibitively expensive to most research labs. While the cost of sequencing is continuing to drop, so is the price of microarrays as the manufactures are actively driving technological advancement to reduce the price of arrays to keep them competitive (Ledford, 2008). As a result, it is expected that the microarrays will continue to be a cheaper and hence, a more affordable option for large scale genetic/genomic analyses.

Considering the comparative advantages and limitations of the two technologies (NGS and microarrays), we envisage that NGS will become a more routine tool for future genetic/genomics research. It would not be surprising if RNA-Seq eventually replace the transcriptome arrays for human research in near future facilitated by fast improvements of the NGS platforms along with their associated bioinformatics tools and with the rapidly plummeting price of sequencing, but the informatics needs are a major issue. Nevertheless, microarray based research will continue to make important contributions for their affordability, ease of application, prevalidated assays, high throughput nature and established mature bioinformatics pipelines. For instance, genotyping arrays will be particularly important in population genetic and conservation genetic research in years to come where analyzing consistent markers across large number of samples is important. Even though the quality and reproducibility of the gene expression and CGH arrays have been a concern, the data generated from genotyping arrays are generally very accurate (accuracy >99.9%) and highly reproducible. Genotyping arrays will be particularly important for genetic improvement of livestock animals through their routine application in GS, where applying routine sequencing will hardly be an alternative. Most likely, a hybrid approach will be used combining the use of genotyping arrays to type common variants and low coverage genome sequencing to capture rare genetic variants. Microarrays are also expected to be a major tool of analysis in clinical diagnostics, a growing field requiring mature technologies with validated assay targets (Ledford, 2008; Shendure, 2008). According to Ledford (2008) improved understanding of the genetic

basis of various diseases will, in fact, create new opportunities for microarrays in clinical diagnostics. Besides, another area where array manufacturers are currently directing much of their efforts is the development of DNA capture arrays which will have a very significant role to play in future life science research as these arrays can be used to isolate candidate regions of the genome (e.g. exomes) for targeted sequencing at a lower cost but at increased coverage leading to higher confidence in variant calls. The usefulness of all types of microarrays will continue to improve with better annotations of the genomes which are improving fast due to many large sequencing DNA and RNA sequencing projects.

CONCLUSIONS

In this review we have seen that microarrays have been invaluable tools for genetic and genomic research in chicken. The arrays that have been commercially developed are important resources for future research not only on chicken but also on other closely related species for comparative genomic analysis. Moreover, the associated databases like EST and cDNA sequences, genetic variants etc. are also major resources for further characterization of the chicken genome. Even though, with the advent of NGS a debate has started over the fate of microarrays, we have argued that microarrays are expected to play important roles in the future research in many fields of life science, particularly in livestock genetics and genomics. Even though sequencing promises many advantages, the future of microarrays will be shaped by many factors like how the two technologies (sequencing and microarrays) develop over the next few years, the nature of the study being undertaken, the type of the organism or system being studied, and the feasibility of applying these technologies in terms of cost, labor, bioinformatics analysis, and scope of the lab in question. The gap between the utility of sequencing and that of microarrays, however, will reduce over time as further improvements of NGS technologies make them more affordable, robust, and less challenging in terms of computing and as the microarrays develop further to be able to incorporate more target regions including rare elements through better annotations of the genomes. We predict that in future, sequencing will be a used more frequently for detection of novel elements in the genome, whereas, the microarrays will be more appropriate for routine screening of large number of individuals on known genomic sites. Most likely, a hybrid approach will be used combining these two technologies. We therefore predict that both these technologies will grow as important genetic and genomic research tools and will become complementary to each other rather than competitors.

LITERATURE CITED

- Abasht B, Lamont SJ. 2007. Genome-wide association analysis reveals cryptic alleles as an important factor in heterosis for fatness in chicken F2 population. *Anim Genet* 38:491–498.
- Aerts J, Megens HJ, Veenendaal T, Ovcharenko I, Crooijmans R, Gordon L, Stubbs L, Groenen M. 2007. Extent of linkage disequilibrium in chicken. *Cytogenet Genome Res* 117:338–345.
- Afrakhte M, Schultheiss TM. 2004. Construction and analysis of a subtracted library and microarray of cDNAs expressed specifically in chicken heart progenitor cells. *Dev Dyn* 230:290–298.
- Alkan C, Coe BP, Eichler EE. 2011. Genome structural variation discovery and genotyping. *Nat Rev Genet* 12:363–376.
- Almmani R, van der Heijden J, Ariyurek Y, Lai Y, Bakker E, van Galen M, Breuning MH, den Dunnen JT. 2011. Experiences with array-based sequence capture; toward clinical applications. *Eur J Hum Genet* 19:50–55.
- Andreescu C, Avendano S, Brown SR, Hassen A, Lamont SJ, Dekkers JCM. 2007. Linkage disequilibrium in related breeding lines of chickens. *Genetics* 177:2161–2169.
- Avendaño S, Watson K, Kranis A. 2010. Genomics in poultry breeding: from Utopia to deliverables. In: *Proceedings of the 9th World Congress on Genetics Applied to Livestock Production*, August 1–6, 2010, Leipzig, Germany. 761 Session 07-01. ISBN 978-3-00-031608-1.
- Bailey MJ, Beremand PD, Hammer R, Bell-Pedersen D, Thomas TL, Cassone VM. 2003. Transcriptional profiling of the chick pineal gland, a photoreceptive circadian oscillator and pacemaker. *Mol Endocrinol* 17:2084–2095.
- Bliss TW, Dohms JE, Emara MG, Keeler CL. 2005. Gene expression profiling of avian macrophage activation. *Vet Immunol Immunopathol* 105:289–299.
- Blum Y, Le Mignon G, Causeur D, Filangi O, Desert C, Demeure O, Le Roy P, Lagarrigue S. 2011. Complex trait subtypes identification using transcriptome profiling reveals an interaction between two QTL affecting adiposity in chicken. *BMC Genomics* 12:567.
- Boardman PE, Sanz-Ezquerro J, Overton IM, Burt DW, Bosch E, Fong WT, Tickle C, Brown WR, Wilson SA, Hubbard SJ. 2002. A comprehensive collection of chicken cDNAs. *Curr Biol* 12:1965–1969.
- Bottje WG, Kong BW, Song JJ, Lee JY, Hargis BM, Lassiter K, Wing T, Hardiman J. 2012. Gene expression in breast muscle associated with feed efficiency in a single male broiler line using a chicken 44K microarray. II. Differentially expressed focus genes. *Poult Sci* 91:2576–2587.
- Brasch-Andersen C, Christiansen L, Tan Q, Haagerup A, Vestbo J, Kruse TA. 2004. Possible gene dosage effect of glutathione-S-transferases on atopic asthma: Using real-time PCR for quantification of GSTM1 and GSTT1 gene copy numbers. *Hum Mutat* 24:208–214.
- Brown WR, Hubbard SJ, Tickle C, Wilson SA. 2003. The chicken as a model for large-scale analysis of vertebrate gene function. *Nat Rev Genet* 4:87–98.
- Bureau C, Hennequet-Antier C, Couty M, Guemene D. 2009. Gene array analysis of adrenal glands in broiler chickens following ACTH treatment. *BMC Genomics* 10:430.
- Burnside J, Neiman P, Tang JS, Basom R, Talbot R, Aronszajn M, Burt D, Delrow J. 2005. Development of a cDNA array for chicken gene expression analysis. *BMC Genomics* 6:13.
- Caldwell RB, Kierzek AM, Arakawa H, Bezzubov Y, Zaim J, Fiedler P, Kutter S, Blagodatski A, Kostovska D, Koter M, Plachy J, Carninci P, Hayashizaki Y, Buerstedde JM. 2005. Full-length cDNAs from chicken bursal lymphocytes to facilitate gene function analysis. *Genome Biol* 6:R6.
- Carson AR, Feuk L, Mohammed M, Scherer SW. 2006. Strategies for the detection of copy number and other structural variants in the human genome. *Hum Genomics* 2:403–414.
- Chiang H-I, Swaggerty CL, Kogut MH, Dowd SE, Li X, Pevzner IY, Zhou H. 2008. Gene expression profiling in chicken heterophils with *Salmonella enteritidis* stimulation using a chicken 44 K Agilent microarray. *BMC Genomics* 9.
- Cloonan N, Forrest AR, Kolle G, Gardiner BB, Faulkner GJ, Brown MK, Taylor DE, Steptoe AL, Wani S, Bethel G, Robertson AJ, Perkins AC, Bruce SJ, Lee CC, Ranade SS, Peckham HE, Manning JM, McKernan KJ, Grimmond SM. 2008. Stem cell transcriptome profiling via massive-scale mRNA sequencing. *Nat Methods* 5:613–619.
- Cogburn LA, Porter TE, Duclos MJ, Simon J, Burgess SC, Zhu JJ, Cheng HH, Dodgson JB, Burnside J. 2007. Functional genomics of the chicken—A model organism. *Poult Sci* 86:2059–2094.
- Cogburn LA, Wang X, Carre W, Rejto L, Porter TE, Aggrey SE, Simon J. 2003. Systems-wide chicken DNA microarrays, gene expression profiling, and discovery of functional genes. *Poult Sci* 82:939–951.
- Davison TE. 2003. The immunologists' debt to the chicken. *Br Poult Sci* 44:6–21.
- de Koning DJ, Cabrera CP, Haley CS. 2007. Genetical genomics: Combining gene expression with marker genotypes in poultry. *Poult Sci* 86:1501–1509.
- Dekkers JCM SC, Fernando RL, Garrick DJ, Lamont SJ, O'Sullivan NP, Fulton JE, Arango J, Settari P, Kranis A, McKay J, Watson KA, Koerhuis A, Preisinger R.

2010. Implementation of genomic selection in egg layer chickens. *J Anim Sci* 87 (E-Suppl. 2/J): iii.
- Elferink MG, Megens HJ, Vereijken A, Hu XX, Crooijmans RPMA, Groenen MAM. 2012. Signatures of selection in the genomes of commercial and non-commercial chicken breeds. *Plos One* 7:e32720.
- Elferink MG, Vallee AA, Jungerius AP, Crooijmans RP, Groenen MA. 2008. Partial duplication of the PRLR and SPEF2 genes at the late feathering locus in chicken. *BMC Genomics* 9:391.
- Elferink MG, van As P, Veenendaal T, Crooijmans RP, Groenen MA. 2010. Regional differences in recombination hotspots between two chicken populations. *BMC Genet* 11:11.
- Ellestad LE, Carre W, Muchow M, Jenkins SA, Wang XF, Cogburn LA, Porter TE. 2006. Gene expression profiling during cellular differentiation in the embryonic pituitary gland using cDNA microarrays. *Physiol Genomics* 25:414–425.
- Fadista J, Nygaard M, Holm LE, Thomsen B, Bendixen C. 2008. A snapshot of CNVs in the pig genome. *Plos One* 3:e3916.
- Fanciulli M, Norsworthy PJ, Petretto E, Dong R, Harper L, Kamesh L, Heward JM, Gough SC, de Smith A, Blakemore AI, Froguel P, Owen CJ, Pearce SH, Teixeira L, Guillemin L, Graham DS, Pusey CD, Cook HT, Vyse TJ, Aitman TJ. 2007. FCGR3B copy number variation is associated with susceptibility to systemic, but not organ-specific, autoimmunity. *Nat Genet* 39:721–723.
- Fontanesi L, Beretti F, Martelli PL, Colombo M, Dall'Olio S, Occidente M, Portolano B, Casadio R, Matassino D, Russo V. 2011. A first comparative map of copy number variations in the sheep genome. *Genomics* 97:158–165.
- Gonzalez E, Kulkarni H, Bolivar H, Mangano A, Sanchez R, Catano G, Nibbs RJ, Freedman BI, Quinones MP, Bamshad MJ, Murthy KK, Rovin BH, Bradley W, Clark RA, Anderson SA, O'Connell R J, Agan BK, Ahuja SS, Bologna R, Sen L, Dolan MJ, Ahuja SK. 2005. The influence of CCL3L1 gene-containing segmental duplications on HIV-1/AIDS susceptibility. *Science* 307:1434–1440.
- Griffin DK, Robertson LB, Tempest HG, Vignal A, Fillon V, Crooijmans RP, Groenen MA, Deryusheva S, Gaginskaya E, Carre W, Waddington D, Talbot R, Volker M, Masabanda JS, Burt DW. 2008. Whole genome comparative studies between chicken and turkey and their implications for avian genome evolution. *BMC Genomics* 9:168.
- Groenen MAM, Megens HJ, Zare Y, Warren WC, Hillier LW, Crooijmans RPMA, Vereijken A, Okimoto R, Muir WM, Cheng HH. 2011. The development and characterization of a 60K SNP chip for chicken. *BMC Genomics* 12:274.
- Groenen MAM, Wahlberg P, Foglio M, Cheng HH, Megens HJ, Crooijmans RPMA, Besnier F, Lathrop M, Muir WM, Wong GKS, Gut I, Andersson L. 2009. A high-density SNP-based linkage map of the chicken genome reveals sequence features correlated with recombination rate. *Genome Res* 19:510–519.
- Gu X, Feng C, Ma L, Song C, Wang Y, Da Y, Li H, Chen K, Ye S, Ge C, Hu X, Li N. 2011. Genome-wide association study of body weight in chicken F2 resource population. *Plos One* 6:e21872.
- Gunderson KL, Steemers FJ, Lee G, Mendoza LG, Chee MS. 2005. A genome-wide scalable SNP genotyping assay using microarray technology. *Nat Genet* 37:549–554.
- Habier D, Fernando RL, Dekkers JCM. 2009. Genomic selection using low-density marker panels. *Genetics* 182:343–353.
- Habig C, Geffers R, Distl O. 2012. Differential gene expression from genome-wide microarray analyses distinguishes Lohmann selected leghorn and Lohmann brown layers. *Plos One* 7:e46787.
- Hackam AS, Bradford RL, Bakhru RN, Shah RM, Farkas R, Zack DJ, Adler R. 2003. Gene discovery in the embryonic chick retina. *Mol Vis* 9:262–276.
- Hardiman G. 2004. Microarray platforms—Comparisons and contrasts. *Pharmacogenomics* 5:487–502.
- Harismendy O, Ng PC, Strausberg RL, Wang X, Stockwell TB, Beeson KY, Schork NJ, Murray SS, Topol EJ, Levy S, Frazer KA. 2009. Evaluation of next generation sequencing platforms for population targeted sequencing studies. *Genome Biol* 10:R32.
- Heidari M, Sarson AJ, Huebner M, Sharif S, Kireev D, Zhou H. 2010. Marek's disease virus-induced immunosuppression: Array analysis of chicken immune response gene expression profiling. *Viral Immunol* 23:309–319.
- Hindorf LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA. 2009. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci USA* 106:9362–9367.
- Hou Y, Liu GE, Bickhart DM, Cardone ME, Wang K, Kim ES, Matukumalli LK, Ventura M, Song J, VanRaden PM, Sonstegard TS, Van Tassell CP. 2011. Genomic characteristics of cattle copy number variations. *BMC Genomics* 12:127.
- Hughes TR, Mao M, Jones AR, Burchard J, Marton MJ, Shannon KW, Lefkowitz SM, Ziman M, Schelter JM, Meyer MR, Kobayashi S, Davis C, Dai H, He YD, Stephanian SB, Cavet G, Walker WL, West A, Coffey E, Shoemaker DD, Stoughton R, Blanchard AP, Friend SH, Linsley PS. 2001. Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer. *Nat Biotechnol* 19:342–347.
- Hurd PJ, Nelson CJ. 2009. Advantages of next-generation sequencing versus the microarray in epigenetic

- research. *Brief Funct Genomic Proteomic* 8:174-183.
- Iafrate AJ, Feuk L, Rivera MN, Listewnik ML, Donahoe PK, Qi Y, Scherer SW, Lee C. 2004. Detection of large-scale variation in the human genome. *Nat Genet* 36:949-951.
- International Chicken Genome Sequencing Consortium. 2004. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* 432:695-716.
- International Chicken Polymorphism Map Consortium. 2004. A genetic variation map for chicken with 2.8 million single-nucleotide polymorphisms. *Nature* 432:717-722.
- Kano R, Konnai S, Onuma M, Ohashi K. 2009. Microarray analysis of host immune responses to marek's disease virus infection in vaccinated chickens. *J Vet Med Sci* 71:603-610.
- Kranis A, Gheyas AA, Boschiero C, Turner F, Yu L, Smith S, Talbot R, Pirani A, Brew F, Kaiser P, Hocking PM, Fife M, Salmon N, Fulton J, Strom TM, Haberer G, Weigend S, Preisinger R, Gholami M, Qanbari S, Simianer H, Watson KA, Woolliams JA, Burt DW. 2013. Development of a high density 600K SNP genotyping array for chicken. *BMC Genomics* 14:59.
- Kong BW, Song JJ, Lee JY, Hargis BM, Wing T, Lassiter K, Bottje W. 2011. Gene expression in breast muscle associated with feed efficiency in a single male broiler line using a chicken 44K oligo microarray. I. Top differentially expressed genes. *Poult Sci* 90:2535-2547.
- Le Bihan-Duval E, Nadaf J, Berri C, Pitel F, Graulet B, Godet E, Leroux SY, Demeure O, Lagarrigue S, Duby C, Cogburn LA, Beaumont CM, Duclos MJ. 2011. Detection of a Cis [corrected] eQTL controlling BCMO1 gene expression leads to the identification of a QTL for chicken breast meat color. *Plos One* 6:e14825.
- Le Mignon G, Desert C, Pitel F, Leroux S, Demeure O, Guernec G, Abasht B, Douaire M, Le Roy P, Lagarrigue S. 2009. Using transcriptome profiling to characterize QTL regions on chicken chromosome 5. *BMC Genomics* 10:575.
- Ledford H. 2008. The death of microarrays? *Nature* 455:847-847.
- Li CL, Wang XL, Li N, Wu CX. 2012. Microarray analysis revealed that immunity-associated genes are primarily regulated by roxarsone in promoting broiler chicken (*Gallus gallus domesticus*) growth. *Poult Sci* 91:3184-3190.
- Li XY, Chiang HI, Zhu J, Dowd SE, Zhou HJ. 2008. Characterization of a newly developed chicken 44K Agilent microarray. *BMC Genomics* 9:60.
- Liu GE, Hou Y, Zhu B, Cardone ME, Jiang L, Cellamare A, Mitra A, Alexander LJ, Coutinho LL, Dell'Aquila ME, Gasbarre LC, Lacalandra G, Li RW, Matukumalli LK, Nonneman D, Regitano LC, Smith TP, Song J, Sonstegard TS, Van Tassell CP, Ventura M, Eichler EE, McDanel TG, Keele JW. 2010. Analysis of copy number variations among diverse cattle breeds. *Genome Res* 20:693-703.
- Maher B. 2008. Personal genomes: The case of the missing heritability. *Nature* 456:18-21.
- Marchini J, Howie B. 2010. Genotype imputation for genome-wide association studies. *Nat Rev Genet* 11:499-511.
- McCarroll SA, Huett A, Kuballa P, Chilewski SD, Landry A, Goyette P, Zody MC, Hall JL, Brant SR, Cho JH, Duerr RH, Silverberg MS, Taylor KD, Rioux JD, Altshuler D, Daly MJ, Xavier RJ. 2008. Deletion polymorphism upstream of IRGM associated with altered IRGM expression and Crohn's disease. *Nat Genet* 40:1107-1112.
- McKinney C, Merriman ME, Chapman PT, Gow PJ, Harrison AA, Highton J, Jones PB, McLean L, O'Donnell JL, Pokorny V, Spellerberg M, Stamp LK, Willis J, Steer S, Merriman TR. 2008. Evidence for an influence of chemokine ligand 3-like 1 (CCL3L1) gene copy number on susceptibility to rheumatoid arthritis. *Ann Rheum Dis* 67:409-413.
- Megens HJ, Crooijmans RPMA, Bastiaansen JWM, Kerstens HHD, Coster A, Jalving R, Vereijken A, Silva P, Muir WM, Cheng HH, Hanotte O, Groenen MAM. 2009. Comparison of linkage disequilibrium and haplotype diversity on macro- and microchromosomes in chicken. *BMC Genet* 10:86.
- Meuwissen THE, Hayes BJ, Goddard ME. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157:1819-1829.
- Miller MB, Tang YW. 2009. Basic concepts of microarrays and potential applications in clinical microbiology. *Clin Microbiol Rev* 22:611-633.
- Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* 5:621-628.
- Muir WM, Wong GK, Zhang Y, Wang J, Groenen MA, Crooijmans RP, Megens HJ, Zhang H, Okimoto R, Vereijken A, Jungerius A, Albers GA, Lawley CT, Delany ME, MacEachern S, Cheng HH. 2008a. Genome-wide assessment of worldwide chicken SNP genetic diversity indicates significant absence of rare alleles in commercial breeds. *Proc Natl Acad Sci USA* 105:17312-17317.
- Muir WM, Wong GK, Zhang Y, Wang J, Groenen MAM, Crooijmans RPMA, Megens HJ, Zhang HM, McKay JC, Mcleod S, Okimoto R, Fulton JE, Settar P, O'sullivan NP, Vereijken A, Jungerius-Rattink A, Albers GAA, Taylor LC, Delany ME, Cheng HH. 2008b. Review of the initial validation and characterization of a 3K chicken SNP array. *Worlds Poult Sci J* 64:219-226.

- Murphy D. 2002. Gene expression studies using microarrays: Principles, problems, and prospects. *Adv Physiol Educ* 26:256–270.
- Neiman PE, Grbic JJ, Polony TS, Kimmel R, Bowers SJ, Delrow J, Beemon KL. 2003. Functional genomic analysis reveals distinct neoplastic phenotypes associated with c-myc mutation in the bursa of Fabricius. *Oncogene* 22:1073–1086.
- Neiman PE, Ruddell A, Jasoni C, Loring G, Thomas SJ, Brandvold KA, Lee RM, Burnside J, Delrow J. 2001. Analysis of gene expression during myc oncogene-induced lymphomagenesis in the bursa of Fabricius. *Proc Natl Acad Sci USA* 98:6378–6383.
- Nielsen R, Paul JS, Albrechtsen A, Song YS. 2011. Genotype and SNP calling from next-generation sequencing data. *Nat Rev Genet* 12:443–451.
- Perez-Enciso M, Ferretti L. 2010. Massive parallel sequencing in animal genetics: Wherefroms and wheretos. *Anim Genet* 41:561–569.
- Pinto D, Darvishi K, Shi XH, Rajan D, Rigler D, Fitzgerald T, Lionel AC, Thiruvahindrapuram B, MacDonald JR, Mills R, Prasad A, Noonan K, Gribble S, Prigmore E, Donahoe PK, Smith RS, Park JH, Hurles ME, Carter NP, Lee C, Scherer SW, Feuk L. 2011. Comprehensive assessment of array-based platforms and calling algorithms for detection of copy number variants. *Nat Biotechnol* 29:U512–U576.
- Preisinger R. 2012. Genome-wide selection in poultry. *Anim Prod Sci* 52:121–125.
- Qanbari S, Hansen M, Weigend S, Preisinger R, Simianer H. 2010. Linkage disequilibrium reveals different demographic history in egg laying chickens. *BMC Genet* 11:103.
- Ramayo-Caldas Y, Castello A, Pena RN, Alves E, Mercade A, Souza CA, Fernandez AI, Perez-Enciso M, Folch JM. 2010. Copy number variation in the porcine genome inferred from a 60 k SNP BeadChip. *BMC Genomics* 11:593.
- Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, Fiegler H, Shapero MH, Carson AR, Chen W, Cho EK, Dallaire S, Freeman JL, Gonzalez JR, Gratacos M, Huang J, Kalaitzopoulos D, Komura D, MacDonald JR, Marshall CR, Mei R, Montgomery L, Nishimura K, Okamura K, Shen F, Somerville MJ, Tchinda J, Valsesia A, Woodwark C, Yang F, Zhang J, Zerjal T, Armengol L, Conrad DF, Estivill X, Tyler-Smith C, Carter NP, Aburatani H, Lee C, Jones KW, Scherer SW, Hurles ME. 2006. Global variation in copy number in the human genome. *Nature* 444:444–454.
- Ren C, Lee MK, Yan B, Ding K, Cox B, Romanov MN, Price JA, Dodgson JB, Zhang HB. 2003. A BAC-based physical map of the chicken genome. *Genome Res* 13:2754–2758.
- Rincon G, Weber KL, Van Eenennaam AL, Golden BL, Medrano JF. 2011. Hot topic: Performance of bovine high-density genotyping platforms in Holsteins and Jerseys. *J Dairy Sci* 94:6116–6121.
- Robb EA, Delany ME. 2012. Case study of sequence capture enrichment technology: Identification of variation underpinning developmental syndromes in an amniote model. *Genes* 3:233–247.
- Roh SW, Abell GC, Kim KH, Nam YD, Bae JW. 2010. Comparing microarrays and next-generation sequencing technologies for microbial ecology research. *Trends Biotechnol* 28:291–299.
- Rubin CJ, Zody MC, Eriksson J, Meadows JRS, Sherwood E, Webster MT, Jiang L, Ingman M, Sharpe T, Ka S, Hallbook F, Besnier F, Carlborg O, Bed'hom B, Tixier-Boichard M, Jensen P, Siegel P, Lindblad-Toh K, Andersson L. 2010. Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature* 464:U587–U145.
- Schmutz J, Grimwood J. 2004. Genomes: fowl sequence. *Nature* 432:679–680.
- Sebat J, Lakshmi B, Troge J, Alexander J, Young J, Lundin P, Maner S, Massa H, Walker M, Chi M, Navin N, Lucito R, Healy J, Hicks J, Ye K, Reiner A, Gilliam TC, Trask B, Patterson N, Zetterberg A, Wigler M. 2004. Large-scale copy number polymorphism in the human genome. *Science* 305:525–528.
- Shendure J. 2008. The beginning of the end for microarrays? *Nat Methods* 5:585–587.
- Shendure J. 2011. Next-generation human genetics. *Genome Biol* 12:408.
- Skinner BM, Robertson LB, Tempest HG, Langley EJ, Ioannou D, Fowler KE, Crooijmans RP, Hall AD, Griffin DK, Volker M. 2009. Comparative genomics in chicken and Pekin duck using FISH mapping and microarray analysis. *BMC Genomics* 10:357.
- Smith J, Sadeyen JR, Paton IR, Hocking PM, Salmon N, Fife M, Nair V, Burt DW, Kaiser P. 2011. Systems analysis of immune responses in Marek's disease virus-infected chickens identifies a gene involved in susceptibility and highlights a possible novel pathogenicity mechanism. *J Virol* 85:11146–11158.
- Smith J, Speed D, Hocking PM, Talbot RT, Degen WGJ, Schijns VEJC, Glass EJ, Burt DW. 2006. Development of a chicken 5 K microarray targeted towards immune function. *BMC Genomics* 7:49.
- Stern CD. 2005. The chick; a great model system becomes even greater. *Dev Cell* 8:9–17.
- Syvanen AC. 2005. Toward genome-wide SNP genotyping. *Nat Genet* 37(Suppl): S5–S10.
- The International HapMap Consortium. 2005. A haplotype map of the human genome. *Nature* 437:1299–1320.
- Theisen A. 2008. Microarray-based comparative genomic hybridization (aCGH). *Nat Educ* 1:1. In: <http://www.nature.com/scitable/topicpage/microarray-based-comparative-genomic-hybridization-acgh-45432>.

- Tsuchihashi Z, Dracopoli NC. 2002. Progress in high throughput SNP genotyping methods. *Pharmacogenomics J* 2:103–110.
- Volker M, Backstrom N, Skinner BM, Langley EJ, Bunzey SK, Ellegren H, Griffin DK. 2010. Copy number variation, chromosome rearrangement, and their association with recombination during avian evolution. *Genome Res* 20:503–511.
- Wang H-B, Li H, Wang Q-G, Zhang X-Y, Wang S-Z, Wang Y-X, Wang X-P. 2007. Profiling of chicken adipose tissue gene expression by genome array. *BMC Genomics* 8:193.
- Wang J, Jiang J, Fu W, Jiang L, Ding X, Liu JF, Zhang Q. 2012. A genome-wide detection of copy number variations using SNP genotyping arrays in swine. *BMC Genomics* 13:273.
- Wang X, Nahashon S, Feaster TK, Bohannon-Stewart A, Adefope N. 2010. An initial map of chromosomal segmental copy number variations in the chicken. *BMC Genomics* 11:351.
- Wang Z, Gerstein M, Snyder M. 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet* 10:57–63.
- Wolc A, Stricker C, Arango J, Settar P, Fulton JE, O'Sullivan NP, Preisinger R, Habier D, Fernando R, Garrick DJ, Lamont SJ, Dekkers JCM. 2011. Breeding value prediction for production traits in layer chickens using pedigree or genomic relationships in a reduced animal model. *Genet Select Evol* 43:5.
- Wong KK, deLeeuw RJ, Dosanjh NS, Kimm LR, Cheng Z, Horsman DE, MacAulay C, Ng RT, Brown CJ, Eichler EE, Lam WL. 2007. A comprehensive analysis of common copy-number variations in the human genome. *Am J Hum Genet* 80:91–104.
- Woo Y, Affourtit J, Daigle S, Viale A, Johnson K, Naggert J, Churchill G. 2004. A comparison of cDNA, oligonucleotide, and Affymetrix GeneChip gene expression microarray platforms. *J Biomol Tech* 15:276–284.
- Wright D, Boije H, Meadows JR, Bed'hom B, Gourichon D, Vieaud A, Tixier-Boichard M, Rubin CJ, Imsland F, Hallbook F, Andersson L. 2009. Copy number variation in intron 1 of SOX5 causes the Pea-comb phenotype in chickens. *PLoS Genet* 5:e1000512.
- Yang TL, Chen XD, Guo Y, Lei SF, Wang JT, Zhou Q, Pan F, Chen Y, Zhang ZX, Dong SS, Xu XH, Yan H, Liu X, Qiu C, Zhu XZ, Chen T, Li M, Zhang H, Zhang L, Drees BM, Hamilton JJ, Papasian CJ, Recker RR, Song XP, Cheng J, Deng HW. 2008. Genome-wide copy-number-variation study identified a susceptibility gene, UGT2B17, for osteoporosis. *Am J Hum Genet* 83:663–674.
- Yauk CL, Berndt ML. 2007. Review of the literature examining the correlation among DNA microarray technologies. *Environ Mol Mutagen* 48:380–394.